

Harvard-Smithsonian Center for Astrophysics

Precision Astronomy Group

MEMORANDUM

To: K.J. Johnston
From: R.D. Reasenberg
Subject: Draft report of the Committee on Data Modeling and Reduction, from Pixels to Product
Date: 20 April 1999
Copy: Committee members

I. Membership.

Tom Greene <tgreene@mail.arc.nasa.gov>
Scott Horner <shorner@usno.navy.mil>
Bill Jefferys <bill@astro.as.utexas.edu>
Dave Monet <dgm@nofs.navy.mil>
Jim Phillips <jphillips@cfa.harvard.edu>
Bob Reasenberg <reasenberg@cfa.harvard.edu>, Chair
Ken Seidelmann <pks@spica.usno.navy.mil>
Sean Urban <seu@pyxis.usno.navy.mil>
David Van Buren <dvb@ipac.caltech.edu>
Richard Vassar <richard.vassar@lmco.com>

II. Overview.

A top level, unified data processing approach is presented for the analysis of the FAME Photometric and Astrometric data. Further, critical aspects of the data processing have been investigated by simulation, and references to the resulting reports are given below.

For the most part, the analysis tool is weighted-least-squares (WLS) fitting, with consideration given to (1) models of the data, (2) statistics of the data, and (3) numerical procedures that keep the processing from being excessively burdensome. However, we see one likely exception. Jefferys is investigating a Bayesian approach to the centroiding problem, which will likely offer advantages for stars near the mission faint limit. The impact of this approach remains to be investigated.

III. Data Flow Diagram and Parameter List.

In the design of large software systems, there are three (four) standard graphical tools used by professionals. (1) State Transition Diagrams are applicable to real-time systems, and will not be addressed further. (2) Structure Charts show the calling relation among modules, and are probably the most familiar in the academic environment. We will eventually need these. (3) Data Flow Diagrams (DFD, which are sometimes extended to include (4) both Data and Control Flow) follow signals (or information or data) through the data-processing system. They are most closely tied to the algorithms used to transform the data.

The DFD (available on the FAME web site) shows an integrated approach to the analysis of astrometric and photometric data. The astrometric stream is shown as an iterative six-step system, where the numbered steps (I, IIa, ...VIa) are the processes that substantively transform the data (cf., selection and sorting.) A short set of explanatory notes accompanies the DFD (also available on the FAME web site). The photometric stream shares the first step (Assign Event Number and Calibrate) with the astrometric stream. There are two additional steps (IIp, IIIp), and use is made of the results of the astrometric analysis.

A parameter list (spread sheet, with printed version available on the FAME web site) is provided as a tool for understanding the data analysis scheme and for estimating the scale of the data storage required to operate the system. It shows that the vast majority of the parameters to be estimated pertain to specific stars. Less than 0.1% of the estimated parameters are not directly related to the stars.

IV. Centroiding Studies.

Centroiding studies have been performed by Chandler and Reasenberg (SAO) and by Monet (USNO, Flagstaff). The latter will be added to the report when a concise summary becomes available from Monet.

The code for the SAO analysis was started by Phillips, who made an investment in numerical techniques that could produce precise diffraction patterns (i.e., photon counts over a fine grid in the detector plane) from a specified broad-band light spectrum in reasonable time. This code also produces the required partial derivatives of the diffraction pattern with respect to spectral and positional parameters, as required for WLS fitting. The code includes two models of stellar spectra, black body and a table traceable to Kurucz (ref.). In a planned refinement, the Gunn-Stryker stars will be added to the code. A recently added feature is the ability to estimate star temperature using the brightness in each of four non-overlapping bands (each 0.125 microns wide) that cover the nominal optical pass band of the instrument (0.4 to 0.9 microns). This can be used both for a starting procedure and as part of the star-parameter estimation scheme.

All studies have been conducted with the old optical specification: focal length = 7.5 m and (each of two) apertures = 25 x 50 cm. Eventually, the key studies will need to be repeated with the new optical design, spin rate, etc. In recent studies, pseudo-data were generated using the Kurucz model and fit using the black-body model. (This use of "crossed models" is clearly a mismatch and intended to show how robust the centroiding is.) The studies assume a $V = 9$ star, an aperture for each look direction of 0.5 by 0.25 m, an optical efficiency of 27% (six reflections at 86% and CCD detection at 65%), a 1.56 sec integration time on the CCD, and a 1% uncertainty in photometry. (Note that the report of the Photometric Filter Subgroup includes an estimate that the photometry of bright stars will be as good as 0.001 to 0.003 mag -- 0.01 mag = 0.92%.) We need to find an uncertainty for the centroided position in the scan direction of no more than 1/700 pixel.

The centroiding is done in batches of seven events, assumed to have been part of the same visit, such that only one star temperature and magnitude need be estimated for the batch. The central results from the studies are:

A. The average (statistical) uncertainty of the centroid (in the scan direction) is $\sigma = 1.0$ to 1.3 milli-pixel, depending on the star temperature. (The "crossed models" has no effect here.)

B. As expected, the "crossed models" yield poor estimates of the physical temperature of the star. In particular, the model error results in a bias large compared to the statistical uncertainty.

C. The "crossed models" do not yield a large bias in the estimate of the centroid. The weighted RMS bias in the centroid, based on the uncertainty in temperature estimated from the color photometry and the (numerically) calculated centroid shift due to a temperature shift, is under 1 milli-pixel. The corresponding mean bias is considerably smaller.

V. Spiral Reduction Studies.

Studies of the Spiral Reduction (Stage III in the DFD, aka Great Circle Reduction) have been performed by Chandler and Reasenberg (SAO) and Germain (USNO, Flagstaff). The latter will be added to the report when additional modeling has been added to the analysis.

The primary function of the spiral reduction is to model the rotation of the instrument, leaving only the estimation of three orientation parameters (e.g., Euler angles) to Stage IV. A measure of the success of the reduction is the rigidity of the rotation model. In particular, one wants to know the uncertainty in the angular separation of one of the look directions at two different times. A statistical measure of this quantity $\sigma(\Delta\phi)$ is used by Chandler and Reasenberg as the figure of merit (TM99-04 is available on the FAME web site). They find that $\sigma(\Delta\phi)/\sigma_0$ (where σ_0 is the single-measurement (centroiding) uncertainty) ranges from 3.0 to 0.03 for the cases considered, when they estimate various "complete" sets of parameters. Key results from their study are:

A. $\sigma(\Delta\phi)$ decreases dramatically as the frequency of attitude correction events (ACE) decreases. Therefore, we must avoid using the Hipparcos approach of firing the attitude control jets several times per spacecraft rotation.

B. Even when the parameter count of the rotation model grows in proportion to the span of uninterrupted rotation, $\sigma(\Delta\phi)$ continues to decrease with increasing number of spacecraft rotations, until it saturates at about 36 rotations. (It is not necessary to operate the instrument at this saturation level.)

C. Although $\sigma(\Delta\phi)$ depends on the basic angle, over a broad range (say ± 50 deg) around the optimum of 100 deg, the variation is slight, except for the bumps at a few "bad angles."

D. Increasing the width (i.e., cross-scan extent) of the field of view offers two advantages. First, there is an increase in the number of observations and a corresponding decrease of $\sigma(\Delta\phi)$ by the "square-root of N." Second, individual stars are observed more times during a visit, which creates a stronger internal connection and results in a further decrease in $\sigma(\Delta\phi)$, comparably important in the domain investigated.

VI. Iteration in the Astrometric Reduction.

[This section needs to be expanded.]

Through the analysis of the data, we will learn about both the observed stars and the observing instrument. A straightforward analysis procedure would put all of the data into a single WLS estimator to yield all of the star parameters and all of the instrument parameters. This will not, by any stretch of the imagination, be possible with available computers. Instead, we will use an iterative (sub-optimal) scheme, such as the one described by Reasenberg and Phillips in SPIE Vol 3356 and implicit in the DFD.

The Spiral Reduction (Stage III in the DFD) will use only a small set of bright stars -- the fiducial stars. It will be followed by a Global Fit (Stage IV) that ties together the spirals. At this stage, it would be possible to produce the catalog (Stage V) by using the global model of the instrument rotation to project all events on the sky and then to fit the position of each object to the projected events. However, to clean up the solution iteratively, the Stage V procedure would be applied to the fiducial stars only. With improved estimates of positions, proper motions, and parallaxes for these fiducial stars, the Stage III and IV analyses would be repeated. Unless this iteration uncovers and precipitates the removal of bad data (blunders) or shows some of the fiducial stars to be unsuitable, it is unlikely that it would *need* to be repeated. However, for other reasons, the iteration is likely to be repeated.

Other aspects of the iterative nature of the analysis include:

A. During the early work with the data, we will learn about the most serious blemishes in the CCD detectors. Later we will learn about additional problems. At some stage, we will apply our knowledge of these problems to the previously processed data.

B. We will detect unseen companions astrometrically in the Stage VI analysis. For the fiducial stars, it will be useful to incorporate this information in a repeat of Stages III and IV. For some stars, it may be useful to to incorporate this information in Stage II.